

УДК 004.75

П.П. Маслянко, А.В. Рябушенко

# КОМПОНЕНТНА МОДЕЛЬ ІНФОРМАЦІЙНО-АНАЛІТИЧНОЇ СИСТЕМИ І ГЕНЕТИЧНИЙ АЛГОРИТМ ФОРМУВАННЯ ОПТИМАЛЬНОГО ПОРТФЕЛЯ АКЦІЙ

## Вступ

Розробка теоретичних положень методології системної інженерії проєктів інформатизації організаційних систем (Орг.С) спирається на фундаментальні основи прикладного системного аналізу і теорії систем [1–6]. Визначення системного критерію поділу проєктів інформатизації на частини не за функціональними, а за загальносистемними ознаками та застосування компонентного процесу розробки проєктів інформатизації дають можливість істотно вдосконалити і систематизувати процеси системного проєктування проєктів інформатизації Орг.С незалежно від типу чи класу інформаційно-комунікаційних систем (ІКС) [7–9]. Застосування цих теоретичних положень до фінансово-інвестиційної сфери діяльності Орг.С дозволяє істотно підняти ефективність бізнесу прискоренням процесу прийняття інвестиційних рішень, зниженням ризиків, збільшенням кількості використаних фінансових інструментів та зменшенням повної вартості володіння інформаційно-комунікаційними технологіями формування оптимального портфеля акцій.

Основною проблемою фінансово-інвестиційної діяльності є задача оптимізації інвестиційного портфеля. Задача оптимізації портфеля акцій постає перед кожним інвестором, який інвестує гроші в цінні папери. Оптимізація портфеля полягає в тому, щоб знайти такий набір акцій, який має оптимальне співвідношення дохідність–ризик, встановлене інвестором.

Гаррі Маркович запропонував аналітичний алгоритм оптимізації портфеля активів [10], за який в 1990 р. одержав Нобелівську премію. Його модель (1) передбачає, що кожний інвестор підбирає собі портфель, виходячи із своєї толерантності до ризику. Дохідність кожної окремої акції вважається стахостичною величиною з нормальним розподілом. Задача оптимізації портфеля зводиться до задачі максимізації корисності інвестора, яка залежить від математичного сподівання дохідності портфеля

і ризику портфеля, що оцінюється за допомогою дисперсії:

$$\begin{aligned} \max_{w_1, \dots, w_n} U(E(r_p), \sigma_p), \\ E(r_p) = \sum_{i=1}^n w_i E(r_i), \\ \sigma_p = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \sigma_{ij}, \\ 0 \leq w_i \leq 1, i = 1, \dots, n, \\ \sum_{i=1}^n w_i = 1, \end{aligned} \quad (1)$$

де  $w_i$  – вага  $i$ -го активу в портфелі;  $U(E(r_p), \sigma_p)$  – функція корисності портфеля, яка відображає особисті побажання інвестора і його толерантність до ризику;  $E(r_p)$  – математичне сподівання дохідності портфеля акцій;  $E(r_i)$  – математичне сподівання дохідності  $i$ -го активу;  $\sigma_p$  – дисперсія дохідності портфеля акцій, яка виступає мірою ризику інвестування;  $\sigma_{ij}$  – коваріація між дохідністю  $i$ -го і  $j$ -го активу;  $n$  – кількість активів, серед яких інвестор має вибір.

Модель Марковича стала значним проривом у моделюванні процесів на фінансовому ринку. Але в алгоритмі Марковича зроблено два припущення: фінансовий ринок має ідеальну ефективність та дохідність активів розподілена нормально з незмінним математичним сподіванням і дисперсією. Ці припущення дуже зручні з математичної точки зору, вони дають можливість звести задачу оптимізації до задачі квадратичного програмування, хоч вони і рідко виконуються на практиці. Будь-яка спроба відійти від названих припущень призводить до задачі оптимізації до комбінаторної задачі, що не розв'язується аналітично.

## Постановка задачі

Мета статті полягає в дослідженні існуючих підходів до моделювання інформаційно-аналітичних систем (ІАС) управління фінансово-інвестиційною діяльністю і розробці компонентної моделі системи формування оптимального портфеля акцій, показі реалізації двох окремих компонентів системи: компонента оптимізації портфеля акцій та компонента про-

гнозування дохідності, дослідженні реалізації в реальних умовах на історичних даних української біржі ПФТС (першої фондової торговельної системи). Об'єктом дослідження є інформаційно-аналітична система управління фінансово-інвестиційною діяльністю, саме її компоненти оптимізації портфеля акцій та прогнозування дохідності. Предмет дослідження – бізнес-процеси оптимізації портфеля акцій та генетичний алгоритм їх моделювання.

### Проблема оптимізації портфеля акцій

Основним методом оптимізації портфеля акцій є модель Марковича. З того часу, коли вона появилась, було зроблено багато спроб розширити її. В публікації [11] вводиться інша міра ризику – середнє абсолютне відхилення (CAB). Ця міра ризику зводить задачу оптимізації портфеля до задачі лінійного програмування, що значно зменшує кількість необхідних розрахунків. Міра CAB є також нечутливою до аутлаєрів, значень, що вибиваються із загального статичного розподілу, та не вимагає нормального розподілу. Проведені дослідження [11] вказують на те, що результати оптимізації портфеля з CAB та дисперсією приблизно еквівалентні (2).

Таким чином, CAB вибирається як міра ризику.

Модель Марковича не вказує, якою саме має бути функція корисності інвестора. В дослідженні була використана найбільш поширена функція корисності – коефіцієнт Шарпа [12]:

$$\begin{aligned} \max_{w_1, \dots, w_n} \frac{E(r_p)}{CAB_p}, \\ E(r_p) = \sum_{i=1}^n w_i E(r_i), \\ CAB_p = \sum_{i=1}^n w_i CAB_i, \\ 0 \leq w_i \leq 1, i = 1, \dots, n, \\ \sum_{i=1}^n w_i = 1, \end{aligned} \quad (2)$$

де  $w_i$  – вага  $i$ -го активу в портфелі;  $\frac{E(r_p)}{CAB_p}$  –

функція корисності портфеля; як функцію корисності вибрано найбільш вживаний показник – коефіцієнт Шарпа [12];  $E(r_p)$  – матема-

тичне сподівання дохідності портфеля акцій;  $E(r_i)$  – математичне сподівання дохідності  $i$ -го активу;  $CAB_p$  – середнє абсолютне відхилення портфеля акцій;  $CAB_i$  – середнє абсолютне відхилення  $i$ -го активу;  $n$  – кількість активів, серед яких інвестор має вибір.

Однак задача (2) все ще має кілька нерозв'язаних проблем. Ця модель не враховує операційних витрат, що виникають під час торгів на біржі, а саме комісійних дилера та різниці цін продажу і купівлі цінних паперів (bid-ask spread). Залишаються також не врахованими можливі зміни статистичного розподілу, які час від часу трапляються на ринку. Дані зміни особливо важко врахувати, оскільки вони роблять значення математичного сподівання змінним у часі:

$$\begin{aligned} \max_{w_1, \dots, w_n, t_1, \dots, t_n} \frac{E(r_p)}{CAB_p}, \\ E(r_p) = \sum_{i=1}^n w_i E(r_i(t_i)), \\ CAB_p = \sum_{i=1}^n w_i CAB_i(r_i(t_i)), \\ \sum_{i=1}^n t_i E(r_i(t_i)) - \sum_{i=1}^n t_i^0 E(r_i^0(t_i^0)) > TC + P_p^0 - P_p, \quad (3) \\ 0 \leq w_i \leq 1, i = 1, \dots, n, \\ 1 \leq t_i \leq T, i = 1, \dots, n, \\ \sum_{i=1}^n w_i = 1, \end{aligned}$$

де  $TC$  – транзакційні витрати;  $P_p^0 - P_p$  – витрати через різницю ціни купівлі і продажу акцій;  $t_i$  – час, на який інвестуються гроші в  $i$ -й актив;  $T$  – термін дії портфеля;  $E(r_i(t_i))$  – математичне сподівання дохідності  $i$ -го активу за період  $t_i$ ;  $CAB(r_i(t_i))$  – середнє абсолютне відхилення  $i$ -го активу за період  $t_i$ .

Наведена постановка задачі оптимізації портфеля акцій є найбільш узагальнювальною (3), але в даному вигляді аналітично нерозв'язною, оскільки функція математичного сподівання дохідності і CAB портфеля акцій залежно від ваги акцій перестають бути неперервними. Ці функції навіть не кусково-неперервні, бо є фактично стохастичними величинами.

### Вибір алгоритму оптимізації

Серед всього різноманіття алгоритмів оптимізації розглянемо основні типи алгоритмів у контексті задачі оптимізації портфеля акцій. Це:

**Гradientні методи** [14]. Щоправда, навряд чи їх можна застосувати для розв'язання цієї задачі, оскільки вони потребують неперервно диференційованої цільової функції оптимізації. Однак, у нашому випадку задача поставлена в дискретному часі, тому функція вважається розривною. Також для функції математичного сподівання доходності не можна припустити жодної із форм кусково-монотонної функції. Тому проблема оптимізації є частково комбінаторною і не може бути представлена як задача оптимізації неперервної функції.

**Вичерпний пошук.** Він потребує для  $N$  акцій щонайменше  $N^{C_N^1 + C_N^2 + \dots + C_N^N} = N^{2^N}$  операцій, тоді для 18 акцій необхідно виконати більше 1e262144 операцій, що є обчислювально складним. А оскільки на даний момент найкращий суперкомп'ютер виконує 1,105e15 операцій за

секунду, то такий алгоритм потребує 1e262123 років для свого завершення.

**Генетичні методи оптимізації.** Ці методи підходять найкраще, бо вони не вимагають ні неперервної диференційованої функції, ні вичерпного пошуку. Основним недоліком їх є те, що вони не гарантують знаходження оптимального значення. Однак вони дозволяють знайти прийнятний результат за відведений час оптимізації.

Таким чином, реалізація постановки задачі з оптимізації портфеля акцій буде розв'язуватися із застосуванням генетичних алгоритмів оптимізації.

### Компонентна модель інформаційно-аналітичної системи

Компонентна модель інформаційно-аналітичної системи управління інвестиційною діяльністю складається з таких компонентів (рис. 1): портал, оптимізації інвестиційного портфеля, управління ризиками, прогнозування, оцінки вартості похідних фінансових інструментів (ПФІ), фінансова інженерія, сек'юритизація, алгоритмічний трейдинг, статистичний арбітраж, арбітраж, інформаційні ресурси, зовнішні обчислювальні ресурси.

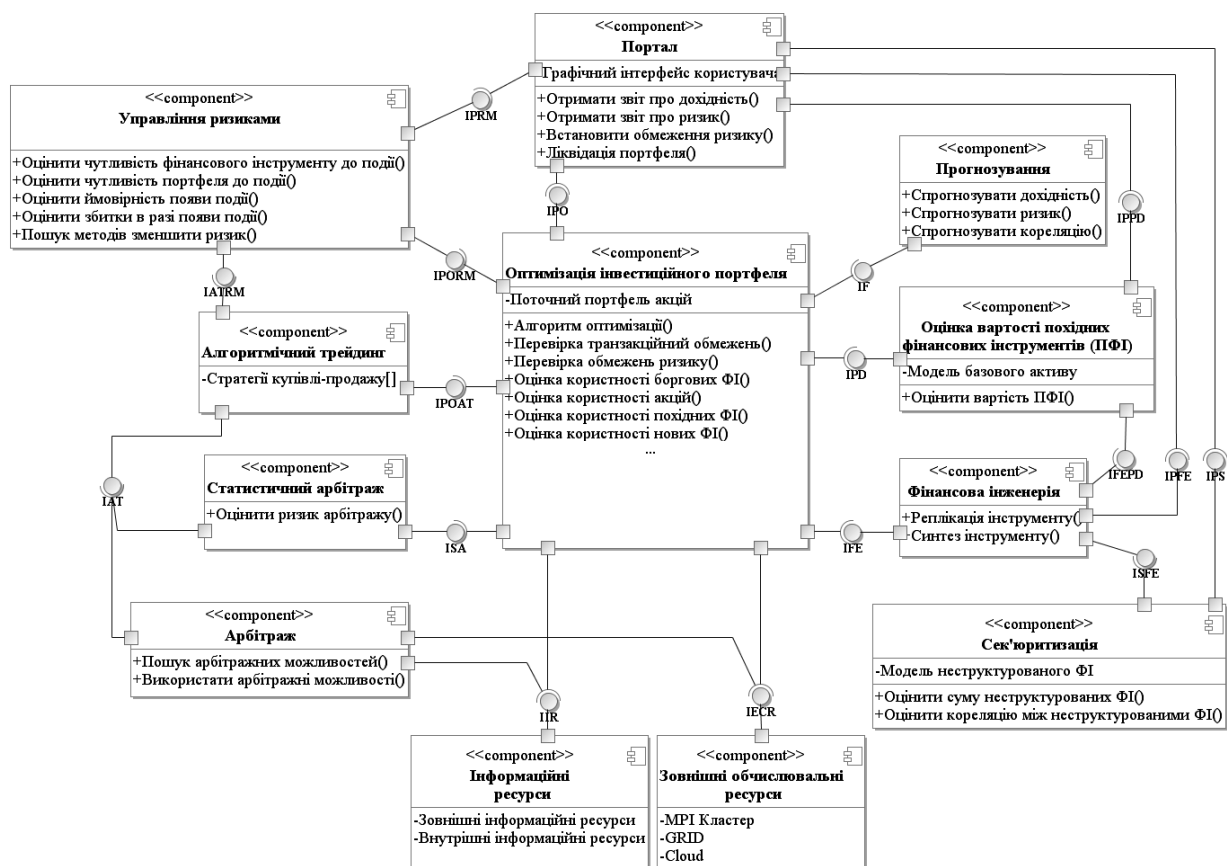


Рис. 1. Модель інформаційно-аналітичної системи і генетичний алгоритм формування оптимального портфеля акцій. Діаграма компонентів у нотації UML

бітразу, статистичного арбітражу, фінансової інженерії, зовнішніх інформаційних ресурсів, зовнішніх обчислювальних ресурсів.

На українському фінансовому ринку, на жаль, немає таких широких можливостей для інвестування, як в розвинутих країнах, тому деякі компоненти автоматизованої системи управління фінансово-інвестиційною діяльністю залишаються незатребуваними. Оцінка вартості похідних фінансових інструментів і фінансової інженерії не використовуються, тому що в Україні нема законодавства про похідні фінансові інструменти. Також нема необхідності застосування компонентів арбітражу, статичного арбітражу та алгоритмічного трейдингу, оскільки жодна з українських бірж не надає інформації про котирування в реальному часі.

Тому найбільш затребуваними в Україні компонентами є компоненти оптимізації портфеля, прогнозування дохідності та управління ризиками. Саме цим компонентам присвячені подальші наші дослідження.

### Специфікація компонентної моделі

Формат специфікації компонентної моделі інформаційно-аналітичної системи управління фінансово-інвестиційною діяльністю формується за основними загальносистемними ознаками – властивостями і операціями [8, 9].

**Назва:** Інформаційно-аналітична система управління фінансово-інвестиційною діяльністю.

#### **Функціональність:**

а) *властивості*: режим функціонування – реальний перебіг часу. Мінімізація часу реакції на зміни котирувань. Вимоги до надійності – 99,999 % живучості системи;

б) *склад*: компонент “Портал” – призначений для надання графічного інтерфейсу користувачу, який дає можливість користуватися всіма доступними послугами інформаційно-аналітичної системи управління фінансово-інвестиційною діяльністю:

- компонент “оптимізація інвестиційного портфеля” – призначений для пошуку такої композиції активів для інвестування, що має оптимальне співвідношення дохідність–ризик;

- компонент “прогнозування” – призначений для прогнозування дохідності, ризику та кореляцій для подальшого використання в оптимізації портфеля;

- компонент “управління ризиками” – призначений для оцінки дохідності портфеля у

всіх можливих випадках і чутливості портфеля до можливих подій у майбутньому. Задача компонента полягає в тому, щоб оцінити можливі збитки в несприятливих випадках та тримати всі можливі збитки у встановлених користувачем рамках;

- компонент “алгоритмічний трейдинг” – призначений для торгівлі цінними паперами на великій швидкості (кілька раз на секунду) при ігноруванні правил оптимального портфеля. Ця стратегія може бути прибутковою, але є дуже ризикованою, тому особливу увагу слід приділяти управлінню ризиками;

- компонент “арбітраж” – призначений для пошуку можливостей без ризику одержати дохідність. На абсолютно ефективному ринку таких можливостей нема, але в реальних умовах вони час від часу з’являються;

- компонент “статистичний арбітраж” – призначений для пошуку можливостей одержати дохідність з ризиком, але таким, який можна надійно оцінити;

- компонент “оцінка вартості похідних фінансових інструментів” – призначений для інвестування у складні фінансові інструменти, що потребують спочатку оцінки справедливої вартості. Це насамперед похідні фінансові інструменти, чия вартість залежить від вартості інших фінансових інструментів;

- компонент “фінансова інженерія” – призначений для підвищення дохідності через створення нових структурних фінансових інструментів, що не продаються і не котуються на ринку, з тих, що вже існують;

- компонент “сек’юритизація” – призначений для створення нових структурних фінансових інструментів через поєднання великої кількості однотипних неструктурних фінансових інструментів в один;

- компонент “зовнішні обчислювальні ресурси” – призначений для забезпечення середовища для проведення обчислювально складних розрахунків, чого потребує вся автоматизована система управління фінансово-інвестиційною діяльністю. Тим компонентом може виступати кластер, GRID, розподілена БД з підтримкою проведення розрахунків на боці БД або cloud;

- компонент “зовнішні інформаційні ресурси” – призначений для забезпечення системи необхідними даними для прийняття інвестиційних рішень. Необхідними даними є котирування з бірж, фінансова звітність компаній,

що котуються (баланс, звіт про прибутки і збитки, звіт про рух коштів), рейтинги кредитоспроможності від рейтингових агентств, макроекономічні індикатори стану економіки країн, курси валют, котирування позабіржового ринку OTC (Over the counter market);

в) *операції*: множина засобів, що забезпечує діяльність інвестора: створення інвестиційного портфеля, задання межі ризику та набору цінних паперів, в які інвестується, отримання звіту про результуючу дохідність та ризик портфеля, можливість розширення функціональності системи новими алгоритмами.

#### Інтерфейси:

- IEIR (Interface External Information Resource) – інтерфейс взаємодії із зовнішніми інформаційними ресурсами. Стандартні інтерфейси передачі фінансових даних: FIX (Financial information exchange), MDDL (the Market Data Definition Language), NewsML, FpML (Financial products Markup Language), RIXML (Research Information Exchange Markup Language), XBRL (Extensible Business Reporting Language), SDMX (Statistical Data and Metadata Exchange) [13];

- IECR (Interface External Computing Resource) – інтерфейс взаємодії із зовнішніми обчислювальними ресурсами. Якості такого інтерфейсу можуть виступати MPI 2.0, MapReduce, SQL:2008;

- IPO (Interface of Portfolio Optimization);
- IPRM (Interface Portal of Risk Management);

- IPORM (Interface Portfolio Optimization of Risk Management);

- IF (Interface Forecasting);

- IPFE (Interface Portal of Financial Engineering);

- IPS (Interface Portal of Securitization);

- ISFE (Interface Securitization of Financial Engineering);

- IPPD (Interface Portal of Pricing Derivatives);

- IFEPD (Interface Financial Engineering of Pricing Derivatives);

- IPD (Interface of Pricing Derivatives);

- IATRM (Interface Algorithmic Trading of Risk Management);

- IPORM (Interface Portfolio Optimization of Risk Management);

- IPOAT (Interface Portfolio Optimization of Algorithmic Trading);

- IAT (Interface of Algorithmic Trading);

- ISA (Interface of Statistical Arbitrage).

#### Призначення

Надати інвестиційним компаніям (хедж-фондам, фондам спільного інвестування, пенсійним фондам, інвестиційним банкам) інтегровану автоматизовану систему управління фінансово-інвестиційною діяльністю, що дасть можливість зменшити витрати на утримання інвестиційних портфелів, зменшити час реакції на зміни, значно прискорити процес прийняття інвестиційного рішення, знайти нові методи отримання додаткової дохідності, а також зробити можливим проведення повномасштабного моделювання інвестиційного портфеля завдяки реалізації інтерфейсу зовнішніх розподілених обчислювальних ресурсів та забезпечити користувача і систему актуальними даними в реальному часі. Специфікація в такому вигляді дає можливість статичного представлення ІАС. Динамічне представлення компонентів генетичної оптимізації портфеля акцій та прогнозування показано на рис. 2.

Діаграма діяльності розкриває призначення, зокрема, компонента генетичної оптимізації портфеля акцій і його зв'язок з компонентом прогнозування та інформаційними ресурсами. Завдання компонента прогнозування – підготувати вхідні дані для подальшої оптимізації. Архітектура компонента має враховувати обчислювальні обмеження всієї системи. Якісні алгоритми прогнозування зазвичай потребують значних обчислювальних ресурсів, причому нові дані котирувань можуть надходити настільки часто, що обчислення нового прогнозу кожного разу, як прийдуть нові дані, буде неможливим навіть з використанням найновітніших суперкомп'ютерів та середовищ GRID. Саме тому компонент прогнозування на діаграмі діяльності заново обчислює прогноз тільки в тому випадку, коли в системі немає актуального прогнозу, який можна використати, чи є вільні обчислювальні ресурси, які можна задіяти.

Після того як актуальний прогноз буде готовий, компонент прогнозування посилає сигнал компоненту оптимізації портфеля, що можна починати оптимізацію. Першим кроком компонент оптимізації портфеля оцінює корисність поточного портфеля без змін, що буде відправною точкою оцінки всіх інших портфелів. Значною перевагою генетичного алгоритму оптимізації є можливість використання паралельних обчислень. Це видно з діаграми діяльності: регіони із стереотипом "parallel" покривають більшу частину алгоритму генетичної оптимізації.

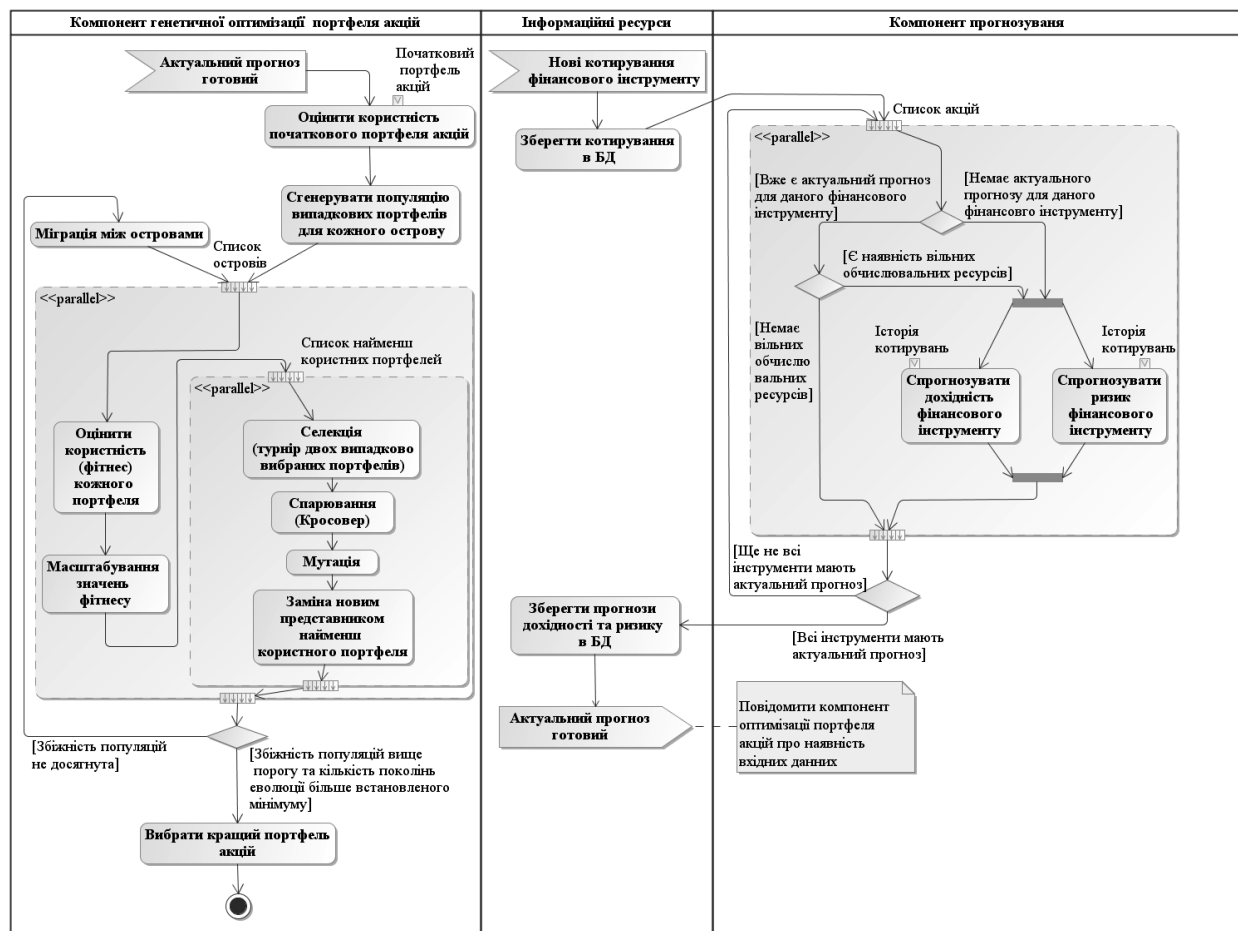


Рис. 2. Модель інформаційно-аналітичної системи і генетичний алгоритм формування оптимального портфеля акцій. Діаграма діяльності в нотації UML

Наступним кроком алгоритм наповнює острови популяціями випадкових портфелів. Далі оцінюється користність кожного портфеля, найгірші портфелі замінюються новими, отриманими в результаті кросоверу. Це забезпечує уточнення знайденого розв'язку. До нових особин, випадковим чином, додається мутація, що дає можливість обстежити нові частини простору потенційних розв'язків задачі. Алгоритм буде повторюватись, доки популяції на всіх островах не досягнуть збіжності. Генетичний алгоритм дуже гнучкий, він дає змогу легко вводити нові обмеження та змінювати існуючі без його істотної зміни.

### Генетичний алгоритм оптимізації портфеля акцій

Генетичний алгоритм вирішує дві проблеми: пошук оптимальної комбінації активів портфеля і пошук найбільш вигідного терміну

інвестування для кожного з активів. Для того щоб розрізнити поняття, які використовуються в цьому дослідженні, термін “актив” використовується в іншому значенні, ніж акції. Актив — це акція з визначеним терміном інвестування, тому акція, яку тримають у портфелі протягом трьох днів є активом, але та сама акція, яку тримають протягом 10 днів, — це вже інший актив. Генетичний алгоритм шукає оптимально комбінацію активів, а не акцій. Він повинен вибрати оптимальну комбінацію серед  $T^N$  можливих активів, де  $T$  — термін дії портфеля, а  $N$  — кількість акцій, які доступні інвестору.

Генетичний алгоритм для оптимізації портфеля розроблений на базі коеволюційного алгоритму [15]. Термінологія, яка використовується в цьому дослідженні, аналогічна використаній у книзі [16].

Далі наведемо схему кодування в нотації генетичних алгоритмів:

- кількість островів дорівнює 4;

- чисельність популяції на кожному острові дорівнює 500 геномів;
- геном — це портфель з 19 генів, активів (18 акцій і гроші);
- ген — це вага активу в портфелі;
- фітнес-функція, найбільш пристосованим геномом є портфель з найвищим коефіцієнтом Шарпа;
- випадковий геном, є портфель з випадковими вагами активів та випадковими термінами інвестування;
- функція кросовера — це зважене арифметичне середнє генів, що зменшує ймовірність появи неможливого портфеля після кросовера;
- ймовірність кросовера 0,9;
- функція вибору представляє турнір двох особин за право розмножуватися;
- ймовірність мутації 0,05;
- топологія міграції — коло.

Ефективність генетичних алгоритмів оптимізації значною мірою залежить від значень параметрів: чисельності населення, кількості островів, ймовірності кросовера, рівня міграції, топології міграції і рівня мутації. Для того щоб відкалібрувати алгоритм і знайти відповідні значення параметрів, які дають можливість одержати оптимальне значення та отримати швидку збіжність, як прогноз дохідності акцій було використано реальні значення. Генетичний алгоритм був реалізований на мові програмування C++ за допомогою бібліотеки GALib, розробленої в MIT (<http://lancet.mit.edu/ga>).

#### **Схема кодування генетичного алгоритму.**

При розробці генетичного алгоритму для конкретної проблеми необхідно прийняти два важливих рішення: зіставлення параметрів проблеми і генів та геномів генетичного алгоритму; визначення цільової функції, яка в нотації генетичного алгоритму називається фітнес-функцією.

**Гени і геноми.** Ген являє собою акцію з двома властивостями: вагою в портфелі ( $w_i$ ) та терміном інвестування ( $t_i$ ). Вага є долею  $i$ -ї акції в портфелі. Така система кодування дозволяє ефективно контролювати обмеження портфеля. Геном — це портфель із 18 акцій та грошей. Сума ваг генів геному повинна дорівнювати одиниці згідно з (3). Геном має додаткові властивості — загальну вартість портфеля

$$P_p = \sum_{i=1}^n p_i q_i, \text{ де } p_i - \text{ціна акції, та } q_i - \text{кількість } i\text{-ї акції.}$$

**Функція ініціалізації геному.** Вона визначена відповідним чином. Кожна вага геному дорівнює випадковому значенню з рівномірного розподілу від нуля до одиниці. Далі, всі ваги діляться на суму ваг, для того щоб сума дорівнювала одиниці. Це обмеження є необхідним за умовою оптимізаційної задачі (3). Термін інвестування вибирається випадково, з рівномірного розподілу від одиниці до  $T$ , де  $T$  — термін дії портфеля.

**Оператор кросоверу.** Це генетичний оператор, який симулює репродукцію відтворення батьків. Кросовер може бути сексуальним або асексуальним. У нашому дослідженні використовується сексуальний кросовер. Він поєднує пару геномів таким чином, що утворюється спадкоємець, який успадковує характеристики своїх батьків. Два геноми мають бути поєднані

так, щоб умова  $\sum_{i=1}^n w_i = 1$  для кожного геному виконувалась.

Ми пропонуємо використати середнє зважене ваг геномів як функцію кросовера геномів. Така функція гарантує виконання необхідної умови і достатньою мірою поєднує характеристики батьків спадкоємця:

$$\lambda w_1 + (1 - \lambda) w_2 = w_n,$$

де  $\lambda$  — випадковий зважувальний коефіцієнт, що лежить у проміжку  $[0,1]$ ;  $w_1$  — вектор ваг акцій геному першого з батьків;  $w_2$  — вектор ваг акцій геному другого з батьків;  $w_n$  — вектор ваг акцій спадкоємця.

**Оператор мутації.** Він випадковим чином змінює деякі гени окремого геному для подальшого вивчення простору розв'язків та забезпечення генетичного різноманіття. Такі випадкові зміни будуть поступово додавати деяких нових особливостей у популяції, які не міг дати кросовер. У більшості випадків поява мутацій має малу ймовірність. Оператор мутації в цьому дослідженні змінює вагу випадкового гена на випадкову величину, що нормально розподілена з математичним сподіванням, яке дорівнює первісній вазі і стандартним відхиленням 0,4; розподіл обрізаний у діапазоні  $[0, 1]$ . Для того щоб компенсувати зміну ваги, інші гени змінюються так, щоб обмеження  $\sum_{i=1}^n w_i = 1$  виконувалось.

Також із ймовірністю 0,5 термін інвестування випадкового гена змінюється на випад-

кову величину, яка рівномірно розподілена на проміжку від 1 до  $T$ , де  $T$  – термін дії портфеля.

**Критерій закінчення алгоритму.** Цей критерій визначає, яке покоління буде останнім. Функція завершення для оптимізації портфеля використовує поріг збіжності популяції на рівні 0,99:

$$\frac{Fitness_{avg}}{Fitness_{best}} \geq 0,99.$$

Однак можливе виникнення проблеми ранньої збіжності популяції в тому випадку, коли початковий портфель близький до оптимального. Для того щоб уникнути даної проблеми, мінімальне число поколінь має становити 50. Також використовуються додаткові заходи від проблеми ранньої збіжності, якщо фітнес найкращого геному менший, ніж у початкового геному, то еволюція продовжується.

**Схема вибору.** Як схема вибору використовується турнір. Два геноми змагаються за право продовжувати покоління в турнірі. Турнір ефективно поєднує дві схеми: схему випадкового вибору (рулетку) і схему впорядкування. Турнір показав значно швидшу і стабільнішу збіжність, ніж методи рулетки та впорядкування.

**Фітнес-функція.** Щоб зробити природний відбір у популяції, кожен геном оцінюється фітнес-функцією. Значення фітнесу вимірює пристосованість геномів (співвідношення дохідності–ризик портфеля), так щоб геноми можна було впорядкувати від найкращого до найгіршого. В задачі оптимізації портфеля фітнес-функція спочатку перевіряє, чи задовольняє портфель обмеження транзакційних витрат. Якщо ні, то фітнес такого портфеля дорівнює  $-\infty$ . В іншому випадку фітнес геному дорівнює коефіцієнту Шарпа портфеля:

$$Fitness_p = \begin{cases} \frac{E(r_p)}{CAB_p}, & \text{якщо} \\ \sum_{i=1}^n t_i E(r_i(t_i)) - \sum_{i=1}^n t_i^0 E(r_i^0(t_i^0)) > TC + P_p^0 - P_p; \\ -\infty, & \text{якщо} \\ \sum_{i=1}^n t_i E(r_i(t_i)) - \sum_{i=1}^n t_i^0 E(r_i^0(t_i^0)) \leq TC + P_p^0 - P_p, \end{cases}$$

де  $\frac{E(r_p)}{CAB_p}$  – коефіцієнт Шарпа портфеля;

$Fitness_p$  – фітнес  $p$ -го геному;  $\sum_{i=1}^n t_i E(r_i(t_i)) -$

$-\sum_{i=1}^n t_i^0 E(r_i^0(t_i^0)) > TC + P_p^0 - P_p$  – транзакційне

обмеження. Прибуток, отриманий в результаті зміни портфеля, має перевищувати витрати на його зміну.

**Масштабування.** Передчасної збіжності також можна уникнути масштабуванням значень фітнесу. Масштабування особливо доцільно використовувати на останніх поколіннях еволюції, коли середнє значення фітнесу наближається до оптимального розв'язку. Це призводить до того, що середній і найкращий геноми мають майже однакову ймовірність бути обраними. В цьому дослідженні використовується сігма-скорочене масштабування, оскільки можливі від'ємні значення фітнесу при оптимізації:

$$F_k = \begin{cases} F - (F_{avg} - c\sigma), & F > 0, \\ 0, & F \leq 0, \end{cases}$$

де  $F$  – початкове значення фітнесу;  $F_k$  – змасштабоване значення фітнесу;  $F_{avg}$  – середнє значення фітнесу всіх геномів популяції;  $c$  – константа для зміни впливу акцій, у випадку оптимізації портфеля акцій  $c = 1$ ;  $\sigma$  – стандартне відхилення фітнесу всіх геномів популяції.

#### Джерела даних, використаних у дослідженні

Україна має ряд фондових бірж. ПФТС – найбільша фондова біржа серед них. Індекс ПФТС отримав статус основного показника стану українського фондового ринку. В нашій статті головну увагу приділено акціям, які входять до індексу ПФТС. У 2007 р. до нього входило 18 акцій. Всі методи оптимізації портфеля і прогнозування перевірялись на всіх 18 акціях.

Дані котирувань і угод від самого першого дня заснування ПФТС були надані компанією КУА “Кінто”. Ця інформація знаходиться у вільному доступі на їх веб-сайті (<http://kinto.com/eng/research/marketupdate/quotes.html>).

Всі компанії в індексі ПФТС мають давню історію котирувань – починаючи з 1997 р. Таким чином, часові ряди є достатньо довги-



ми, щоб захопити довгострокову пам'ять. Тим не менше, в період з 1997 по 2004 рр. ринок ПФТС був досить молодим і неліквідним. В той час траплялися дні, що проходили без торгів, це призводило до втрати даних котирувань за окремі дні торгів. Такі дні були інтерпольовані кубічним сплайном. Тоді траплялися й стрибки цін з близького до нуля значень, що призводило до нескінченної дохідності. Такі стрибки також були інтерпольовані кубічним сплайном, для того щоб уникнути проблем з оцінкою статистичних моделей прогнозування.

### Методика проведення апробації генетичного алгоритму на ПФТС

За критерій оцінки алгоритму оптимізації портфеля прийнято річну дохідність на ПФТС за 2007 р. Цей критерій має ряд переваг: чітку економічну інтерпретацію, можливість порівняти з іншими простими інвестиційними стратегіями, такими, як купівля акцій у рівних частках чи купівля акцій, пропорційних до капіталізації компаній. До недоліків же можна віднести залежність критерію від часу та місця, а також неможливість екстраполювати результати на інші біржі і роки.

Моделювання здійснювалось за таким алгоритмом: у кожний з 246 торгових днів за 2007 р. слід повторити відповідні кроки:

- побудувати прогноз ціни купівлі і продажу для кожної акції із застосуванням алгоритму прогнозування, використовуючи історії котирувань, доступних на той день;
- розв'язати задачу оптимізації портфеля акцій генетичним алгоритмом оптимізації;
- перебалансувати портфель, виходячи з рекомендацій алгоритму оптимізації;
- розпочати новий торговий день;
- підрахувати дохідність портфеля за день;
- перейти до кроку 1.

У дослідженні було приділено багато уваги реалістичності умов моделювання. Для цього враховувались такі фактори:

- транзакційні витрати, комісійні дилеру за обслуговування в розмірі 0,5 % з транзакції;
- різниця цін купівлі і продажу;
- в розрахунках брали участь 18 акцій, що входили в індекс ПФТС в 2007 р., плюс гроші як один з видів фінансових активів;
- гроші розглядались як безризиковий актив з від'ємною дохідністю, що пояснюється інфляцією;

- інфляція за 2007 р. становила 16,6 %, тому денна дохідність грошей була взята на рівні  $-16,6\%/246 = -0,067\%$ ;

- короткі продажі не допускаються згідно з правилами ПФТС, тому вони були вилучені з оптимізації;

- початковий портфель на перший день 2007 р. становив 1 млн доларів.

### Результати апробації генетичного алгоритму на ПФТС

Результати симуляції за 2007 р. подані в таблиці. У другій колонці містяться результати, які враховують транзакційні витрати, але без врахування різниці цін купівлі і продажу. Як ціна була використана середньозважена ціна котирування. У другій колонці подано результати, що враховують як транзакційні витрати, так і різницю цін купівлі і продажу. Для порівняння в таблиці наведено дохідність пасивної інвестиційної стратегії, що є найбільш популярною серед інвесторів. Механізм стратегії полягає в тому, що акції купуються в тих самих частках, що й в індексі. Пасивна інвестиційна стратегія в 2007 р. принесла дохідність у 135,8 % (із врахуванням витрат 134 %).

**Таблиця.** Результати апробації алгоритмів оптимізації портфеля акцій порівняно з пасивною інвестиційною стратегією

Метод прогнозування	Річна дохідність на ринку ПФТС за 2007 рік	
	Із врахуванням транзакційних витрат, %	Із врахуванням транзакцій та різниці цін купівлі та продажу, %
Пасивна інвестиційна стратегія без прогнозування	135,8	134
Генетичний алгоритм з ідеальним прогнозом	1913–1982	1734–1760
Генетичний алгоритм з мультифрактальною моделлю	187,2–190,9	159,8–162

З третього по четвертий рядки таблиці міститься дохідність, отримана з використанням генетичного алгоритму. Генетичний алгоритм — це стохастичний алгоритм, тому в результатах є невелика розбіжність. У третьому рядку наведено результати генетичної оптимізації у випад-

ку, коли як прогноз були використані справжні значення. В цьому випадку було отримано дохідність, яка перевищувала 1900 % (1700 % із врахуванням витрат). Ця цифра наочно підтверджує ефективність генетичного алгоритму оптимізації портфеля акцій. Але в реальних умовах майбутні значення невідомі і мають бути спрогнозовані статистичними моделями.

У четвертому рядку таблиці містяться результати генетичної оптимізації з використанням мультифрактальної моделі прогнозування [17]. Дохідність у середньому становить 189 % (161 % із врахуванням витрат). Вона значно перевищила пасивну інвестиційну стратегію, тому можна зробити висновок щодо економічної доцільності впровадження в інвестиційних установах методу генетичної оптимізації портфеля акцій з мультифрактальною моделлю прогнозування.

Згідно з [investfunds.com.ua](http://investfunds.com.ua), тільки 6 фондів з 41 показали дохідність, вищу за 159,8 % [18]. Це підтверджує високу ефективність запропонованого генетичного алгоритму для оптимізації портфеля акцій.

### Висновки

Застосування компонентної моделі до створення інформаційно-аналітичних систем формування портфеля акцій дозволяє формалізувати архітектуру і топологію сутностей інформатизації бізнес-процесів у формуванні портфеля акцій.

Для реалізації компонента оптимізації портфеля пропонується застосування генетич-

ного алгоритму, що дає можливість знайти розв'язання частково комбінаторної задачі. Це рішення дозволяє усунути обмеження класичної моделі Марковича та істотно вдосконалити її. Запропонована узагальнена модель сприяє оптимізації портфеля акцій в умовах, коли статистичний розподіл дохідності акцій невідомий. Дана модель також враховує комісійні витрати, які неминуче присутні на будь-якому фінансовому ринку.

Практична цінність даного дослідження полягає в тому, що запропонований алгоритм оптимізації портфеля акцій піднімає його дохідність і багаторазово прискорює процес ухвалення рішення на фондовому ринку, звільняє аналітичні відділи інвестиційних компаній від тривалої і складної рутинної роботи. Завдяки можливості розпаралелювання генетичного алгоритму та, як наслідок, можливості ефективного використання багатопроцесорних і кластерних комп'ютерних систем даний алгоритм відкриває нові можливості щодо створення наддиверсифікованих портфелів з великої кількості акцій. Дані портфелі акцій дозволяють знизити ризик без зниження прибутковості. Раніше створення таких великих портфелів було неможливим через занадто велику кількість операцій, що виконувались вручну. Також було проведено тестування алгоритму на реальних даних за 2007 р., отриманих у ПФТС (перша фондова торговельна система). Тестування підтвердило доцільність провадження генетичного алгоритму оптимізації портфеля акцій в інвестиційних компаніях.

П.П. Маслянюк, А.В. Рябушенко

КОМПОНЕНТНАЯ МОДЕЛЬ ИНФОРМАЦИОННО-АНАЛИТИЧЕСКОЙ СИСТЕМЫ И ГЕНЕТИЧЕСКИЙ АЛГОРИТМ ФОРМИРОВАНИЯ ОПТИМАЛЬНОГО ПОРТФЕЛЯ АКЦИЙ

Показаны результаты моделирования информационно-аналитической системы управления финансово-инвестиционной деятельностью и разработка компонентной модели системы формирования оптимального портфеля акций. Продемонстрирована реализация двух отдельных компонентов системы: компонента оптимизации портфеля акций и компонента прогнозирования до-

P.P. Maslyanko, A.V. Riabushenko

THE COMPONENT MODEL OF INFORMATION ANALYSIS SYSTEM AND GENETIC ALGORITHM OF OPTIMAL EQUITIES PORTFOLIO FORMATION

The present paper shows the modeling results of information analysis control system of financial investment and development of component model of optimal equity portfolio system. We demonstrate the implementation of two system components: optimization of equity portfolio and return forecasting. Moreover, we test the models on the historical data of Ukrainian stock exchange (PFTS).

ходности. Апробация моделей проведена на исторических данных украинской Первой фондовой торговой системы.

Объект исследования – информационно-аналитическая система управления финансово-инвестиционной деятельностью: ее компоненты оптимизации портфеля акций и прогнозирования доходности. Предмет исследования – бизнес-процессы оптимизации портфеля акций и генетический алгоритм их моделирования.

Our objective is to study the information analysis control system of financial investment: its components of equity portfolio optimization and returns forecasting. Our research subject is the business processes of equity portfolio optimization and the genetic algorithm of their modeling.

1. Згуровський М.З., Панкратова Н.Д. Системний аналіз: проблеми, методологія, застосування. – К.: Наук. думка, 2005. – 744 с.
2. Згуровський М.З., Родіонов М.К., Жилієв І.Б. Розвиток інформаційного суспільства в Україні. – К.: НТУУ "КПІ", 2006. – 544 с.
3. Ericsson H.-E., Penker M. Business Modeling with UML: Business Patterns at work. – Wiley Computer Publishing, 2000. – 350 p.
4. Unified Modeling Language Specification, Version 2.0. Object Management Group, Framingham, Mass., 2004. [www.omg.org](http://www.omg.org).
5. Rumbaugh J., Jacobson I., Booch G. The Unified Modeling Language Reference Manual, Second Edition 2004, by Pearson Education, Inc.
6. Маслянюк П.П. Концепція інформатизації корпоративних структур // Наукові вісті НТУУ "КПІ". – 2001. – № 3. – С. 45–50.
7. Маслянюк П.П. Основні положення методологій системного проектування інформаційно-комунікаційних систем // Там же. – 2007. – № 6. – С. 54–60.
8. Маслянюк П.П. Системне проектування процесів інформатизації // Там же. – 2008. – № 1. – С. 28–36.
9. Маслянюк П.П. Компонентні процеси системного проектування інформаційно-комунікаційних систем // Там же. – № 2. – С.112–121.
10. Markowitz H. Portfolio Selection // J. of Finance. – 1952. – N 7. – P. 77–91.
11. Feinstein C., Thapa M. A reformulation of a mean-absolute deviation portfolio optimization model // Management Science. – 1993. – N 39. – P. 1552–1553.
12. Sharpe W.F. Mutual fund performance measurement // J. of Bussiness. – 1966. – P. 118–138.
13. Martin S. MDDL and the quest for a market data. – Elsevier, 2007. – 320 p.
14. Банди Б. Методы оптимизации: Вводный курс. – М.: Радио и связь, 1988. – 128 с.
15. Koza J.R. Genetic programming: On the Programming of Computers by Mean of Natural Selection. – Cambridge: Cambridge Press, 1992.
16. Chambers D. The practical handbook of genetic algorithms, Applications. Chapman & Hall/CRC, 2001.
17. Lux T., Kaizoji T. Forecasting volatility and volume in the Tokyo Stock Market: Long Memory, Fractality and Regime switching // J. of Economic Dynamics and Control. – Elsevier, 2007. – 6, N 31. – P. 1808–1843.
18. <http://investfunds.com.ua/funds/rate.phtml?period=year>

Рекомендована Радою  
факультету прикладної математики  
НТУУ "КПІ"

Надійшла до редакції  
30 січня 2009 року